

RESEARCH ARTICLE

stelfi: An R package for fitting Hawkes and log-Gaussian Cox point process models

Charlotte M. Jones-Todd  | Alec B. M. van Helsdingen

Department of Statistics, University of Auckland, Auckland, New Zealand

Correspondence

Charlotte M. Jones-Todd, Department of Statistics, University of Auckland, Auckland, New Zealand.

Email: c.jonestodd@auckland.ac.nz**Funding information**

Asian Office of Aerospace Research and Development, Grant/Award Number: FA2386-21-1-4028; Royal Society Te Apārangi, Grant/Award Number: UOA 3723517

Abstract

Modelling spatial and temporal patterns in ecology is imperative to understand the complex processes inherent in ecological phenomena. Log-Gaussian Cox processes are a popular choice among ecologists to describe the spatiotemporal distribution of point-referenced data. In addition, point pattern models where events instigate others nearby (i.e., self-exciting behaviour) are becoming increasingly popular to infer the contagious nature of events (e.g., animal sightings). While there are existing R packages that facilitate fitting spatiotemporal point processes and, separately, self-exciting models, none incorporate both. We present an R package, *stelfi*, that fits spatiotemporal self-exciting and log-Gaussian Cox process models using Template Model Builder through a range of custom-written C++ templates. We illustrate the use of *stelfi*'s functions fitting models to Sasquatch (bigfoot) sightings data within the USA. The structure of these data is typical of many seen in ecology studies. We show, from a temporal Hawkes process to a spatiotemporal self-exciting model, how the models offered by the package enable additional insights into the temporal and spatial progression of point pattern data. We present extensions to these well-known models that include spatiotemporal self-excitation and joint likelihood models, which are better suited to capture the complex mechanisms inherent in many ecological data. The package *stelfi* offers user-friendly functionality, is open source, and is available from CRAN. It offers the implementation of complex spatiotemporal point process models in R for applications even beyond the field of ecology.

KEYWORDS

Hawkes process, log-Gaussian Cox process, point-referenced data, self-exciting, spatiotemporal, Template Model Builder

TAXONOMY CLASSIFICATION

Spatial ecology

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Point patterns are data that describe the locations of objects or events in space and time. For example, the location of a particular species, its nest or the timestamp of its sighting are all point pattern data. A point pattern is a random variable, a single realisation of some assumed point process that describes the characteristics of the observed spatiotemporal distribution (Baddeley et al., 2015; Diggle, 2013; Illian et al., 2008; Ripley, 1981).

Point pattern data are widely seen in ecology (Ben-Said, 2021; Ripley, 1987; Velázquez et al., 2016; Wiegand & Moloney, 2014). With rich spatial and temporal data becoming increasingly available ecologists require progressively sophisticated point process models to capture the structure and dependencies inherent in such data. Summary statistics (e.g., Ripley's K-function) are only able to give ecologists a synoptic description of the data (e.g., the degree of deviation from a homogeneous Poisson process). In reality, an ecologist's interest lies in inferring the effect of environmental covariates and/or characteristics of the points (i.e., marks) on the spatiotemporal distribution of the ecological objects (Velázquez et al., 2016; Wiegand & Moloney, 2014).

Modelling spatial point pattern data as a log-Gaussian Cox process (Cox & Isham, 1980; Møller et al., 1998) is a popular choice for many ecological applications (Illian et al., 2012; Serra et al., 2014; Simpson et al., 2016; Soriano-Redondo et al., 2019). However, in many situations, the occurrence times of events are self-exciting (i.e., the occurrence of an event instigates another). One well-known model describing this phenomenon is a Hawkes process (Hawkes, 1971a, 1971b); these models, and their extensions, are often seen in seismology (Ogata, 1988), criminology (Park et al., 2021; Zhuang & Mateu, 2019) and finance (Bacry et al., 2015; Hawkes, 2018). Recently, however, they have begun to see some use in ecology (Gupta et al., 2018; Nakagawa et al., 2019).

Additionally, the characteristics (i.e., marks) of the points may depend on the locations of the objects. For example, certain species of shrub may be more likely to cluster together (higher point intensity). This can be formulated similarly to a preferential sampling model where the log-Gaussian Cox model assumed for the point pattern (the locations) is modelled jointly alongside the response (i.e., mark) (Diggle & Su, 2010).

We introduce the R package *stelfi* (Jones-Todd & van Helsdingen, 2023), available from the Comprehensive R Archive Network (CRAN), which fits spatiotemporal point process models using Template Model Builder (Kristensen et al., 2016). Specifically, *stelfi* allows users to fit temporal self-exciting Hawkes models (Hawkes, 1971a, 1971b), spatial and spatiotemporal log-Gaussian Cox process models (Cox & Isham, 1980) and self-exciting spatiotemporal models.

The most widely known and user-friendly R packages that fit log-Gaussian Cox processes include *spatstat* (Baddeley et al., 2015), *lgcp* (Taylor et al., 2013), and *inlabru* (Bachl et al., 2019) a wrapper for *INLA* (Lindgren & Rue, 2015). The package *spatstat* offers a comprehensive suite of methods to

summarise and model point pattern data. When fitting these models, the user can choose either a minimum contrast, second-order composite likelihood or a Palm likelihood approach. The packages *lgcp* and *inlabru* fit models using a Bayesian framework via Markov-chain Monte Carlo (Robert et al., 1999) and Laplace approximation, respectively.

Just over a decade ago Rue et al. (2009) developed the Integrated Nested Laplace Approximation approach that uses Laplace approximation techniques to fit latent Gaussian models. A few years later, the Stochastic Partial Differential Equation approach was established by Lindgren et al. (2011). This methodology approximates the Gaussian field by a Markovian equivalent, a Gaussian Markov Random Field. Although *INLA* (Lindgren & Rue, 2015) does so, it is not necessary to use this Bayesian approach if you wish to make use of this framework. We use Template Model Builder (Kristensen et al., 2016); here, again, the Laplace approximation is used to approximate the integration across the random effects (Skaug & Fournier, 2006).

There are a number of R packages that fit temporal Hawkes process models: *emhawkes* (Lee, 2021) where estimation is based on the maximum likelihood method introduced by Ozaki (1979); *hawkes-bow* (Cheysson, 2021), which fits a Hawkes process to discrete data by minimising the Whittle contrast; *hawkes* (Zaatour, 2014) that only allows users to evaluate the Hawkes likelihood for their own optimisation technique *bayesianETAS* (Ross, 2017), and most recently *ETAS.inlabru* (Naylor & Serafini, 2023; Serafini et al., 2023), allow users to fit epidemic-type aftershock sequence models (a special case of a Hawkes process typically used to model the evolution of seismicity over time and space; Ogata, 1988) using Bayesian estimation techniques.

In contrast to the packages mentioned above, the R package *stelfi* uses custom Template Model Builder (Kristensen et al., 2016) templates written in C++ to fit spatiotemporal Hawkes and log-Gaussian Cox models. Where appropriate, the Laplace approximation is used to approximate the integration over the random effects (Skaug & Fournier, 2006). We present extensions to the well-known Hawkes and log-Gaussian Cox process, including spatiotemporal self-excitation and joint likelihood models, which are better suited to capture the complex mechanisms inherent in many ecological data.

The package *stelfi* is available on CRAN, with the development version on GitHub <https://github.com/cmjt/stelfi>. The novelty of the package is twofold (1) custom-written C++ templates take advantage of the R package *TMB* (Kristensen et al., 2016), fitting models via maximum likelihood; (2) it facilitates the fitting of temporal and spatiotemporal self-exciting point process models.

In the sections below we outline the range of spatial, temporal, and spatiotemporal point process models offered by *stelfi* and move on to illustrate their use by modelling Sasquatch sighting data, collected by the Bigfoot Field Researchers Organization (BFRO) (BFRO, 1995) and collated by Renner (2021). These data contain the temporal and spatial locations of claimed sightings of Sasquatch across the contiguous USA ($A \sim 7,771,155 \text{ km}^2$) from 2000 to 2006. Although the data pertain to sightings of the likely fictional creature,

colloquially known as Bigfoot, they are legitimate and share a structure common to many ecological examples (i.e., spatial and temporal coordinates). These data are shipped with the package `stelfi` as the object `sasquatch`.

2 | METHODS AND FEATURES

The core functions of `stelfi` allow users to fit (1) temporal Hawkes point processes, with the extension of a user-defined temporally varying background function and multiple correlated streams; (2) spatial or spatiotemporal log-Gaussian Cox processes, with the extension of jointly modelling relevant marks; (3) self-exciting spatiotemporal models where the spatial self-excitement may be either time-dependent or not.

Before we outline the functionality offered by `stelfi` we will first list the acronyms we will use, for brevity, from here on. Where necessary we also give a brief description of the context we assume, please see the references for the full definitions. All acronyms are commonly used in their respective literature.

- GRF, Gaussian Random Field, a collection of (typically 2D) Gaussian distributed random variables.
- GMRF, Gaussian Markov Random Field, the Markovian equivalent of a GRF (i.e., where the Markov property of conditional independence holds).
- LGCP, Log-Gaussian Cox Process, a commonly used statistical model for clustered point pattern data, which includes a latent GRF (Cox & Isham, 1980).
- INLA, Integrated Nested Laplace Approximation, a Bayesian method for estimating parameters of latent Gaussian model (e.g., the LGCP) (Lindgren & Rue, 2015).
- SPDE, Stochastic Partial Differential Equation, an equation whose solution under certain conditions is the GMRF approximation to the assumed GRF (Lindgren et al., 2011).
- TMB, Template Model Builder, an R package for fitting statistical latent variable models (Kristensen et al., 2016).

2.1 | The Hawkes processes

Hawkes processes (Hawkes, 1971a, 1971b) are self-exciting temporal processes, where the occurrence of one event increases the probability of events in the near future. This makes them a particularly useful tool in describing clustering and interactions between events (i.e., where one event might induce the occurrence of another in near proximity). A Hawkes process has event (e.g., sighting of a species) times $0 < \tau_1 < \tau_2 < \dots < T$ where the first event occurs at τ_1 . After each event the intensity immediately increases, meaning that the occurrence of one event 'excites' another by some degree. Many ecological or environmental data are the result of some contagious effect. For example, the (random) sighting of a species might induce a sighting in the near future only because one is preinclined

to sight it. Modelling these data with a self-exciting process enables the degree of contagion (i.e., self-excitement) to be inferred.

Figure 1 plot (a) shows a point pattern where the observed event occurrences are shown by the points along the x-axis. The Hawkes intensity is given by the solid line and can be thought of as akin to the chance of observing an event at any time t , which increases immediately after an event occurs and decays exponentially over time if no event is observed for some period.

For current time t the conditional intensity function, $\lambda(t)$ in Figure 1 plot (a), is given by

$$\lambda(t) = \mu + \alpha \sum_{i: \tau_i < t} \exp(-\beta^*(t - \tau_i)).$$

Here μ is the background Poisson rate of the process. The term $\alpha \sum_{i: \tau_i < t} \exp(-\beta(t - \tau_i))$ describes the historic temporal dependence (i.e., for times $\tau_i < t$, $i = 1, \dots, T$). The parameter α is the increase in intensity immediately after the occurrence of an event, and $\beta > 0$ controls the exponential decay of the intensity if no event has occurred. An extension of this includes mark information and is given by

$$\lambda(t; m(t)) = \mu + \alpha \sum_{i: \tau_i < t} m(\tau_i) \exp(-\beta(t - \tau_i)). \quad (1)$$

Here $m(t)$ is the temporal mark that multiplies the self-exciting component of $\lambda(t; m(t))$. The package `stelfi` implements additional extensions where (1) the background rate can vary in time where μ in Equation (1) becomes $\mu(t)$ according to some user-defined function and (2) the process can be self-inhibitive or self-correcting (i.e., $\alpha < 0$). Examples of these extensions are given in the package's online gitbook <https://cmjt.github.io/stelfi/>.

2.2 | A multivariate Hawkes process

A multivariate Hawkes process has multiple types of events (streams or threads) where the event occurrence of any type of event is influenced by all past events (i.e., event of any type). Consider the sightings example given in the section above, but now we have two sympatric (co-occurring) species where after the (random) sighting of any one species the chance of observing either species immediately increases. This is an example of a bivariate Hawkes process where both self-excitement (i.e., the chance of sighting a species increases immediately after observing it) and cross-excitement (i.e., the chance of sighting a species increases immediately after observing a sympatric species) may exist. Fitting a multivariate Hawkes process enables the degree of the excitement within (self-excitement) and between (cross-excitement) event types to be inferred.

Figure 1 plot (b) shows a bivariate Hawkes process. The intensity of any stream, $\lambda_i(t)$, increases immediately after any event occurs and decays exponentially over time if no event is observed for some period. The influence of one stream on another may differ. For example, in Figure 1 plot (b) the point pattern is simulated

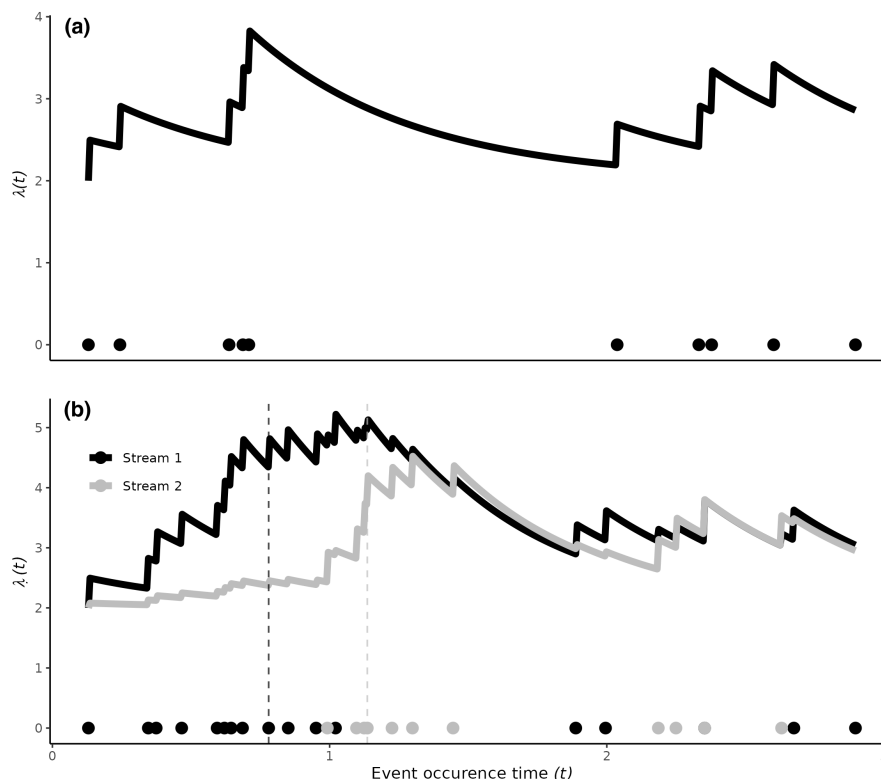


FIGURE 1 Plot (a) Realisation of a Hawkes process where the observed event occurrences are shown by the points. The Hawkes intensity, $\lambda(t)$, is given by the solid line, increases immediately after an event occurs and decays exponentially over time if no event is observed for some period. Plot (b) A realisation of a bivariate Hawkes process where the observed event occurrences from two streams are shown by the coloured points. The Hawkes intensities, $\lambda_{\text{Stream1}}(t)$ and $\lambda_{\text{Stream2}}(t)$, are given by the solid lines. Each intensity exhibits within- and between-stream excitement. The black dashed vertical line shows that a Stream 1 event causes a jump in $\lambda_{\text{Stream1}}(t)$ as well as, to a smaller degree, $\lambda_{\text{Stream2}}(t)$; The grey dashed vertical line shows that a Stream 2 event again causes a jump both intensities. This bivariate Hawkes was simulated with greater within-stream excitement (self-excitement) than between-stream excitement (cross-excitement).

from a multivariate Hawkes process where the background rate, the within-stream excitement, and the exponential decay for each are the same, however, the within-stream influence is greater than the between-stream influence. In plot (b) in Figure 1 this cross-excitation effect is evident; for example, at the dashed horizontal lines, both intensities jump irrespective of the observed stream event.

Formally, the conditional intensity for the j th ($j = 1, \dots, N$) stream is given by

$$\lambda(t)^j = \mu_j + \sum_{k=1}^N \sum_{i: \tau_i < t} \alpha_{jk} e^{(-\beta_j(t-\tau_i))}, \quad (2)$$

where $j, k \in (1, \dots, N)$. Here, α_{jk} is the excitement caused by the k th stream on the j th. Therefore, α is an $N \times N$ matrix where the diagonals represent the within-stream excitement, and the off-diagonals represent the excitement between streams.

2.3 | The log-Gaussian Cox process

A LGCP (Møller et al., 1998) is a commonly used model for spatially clustered point pattern data. It is often called a hierarchical, or doubly stochastic, process as it is an inhomogeneous Poisson

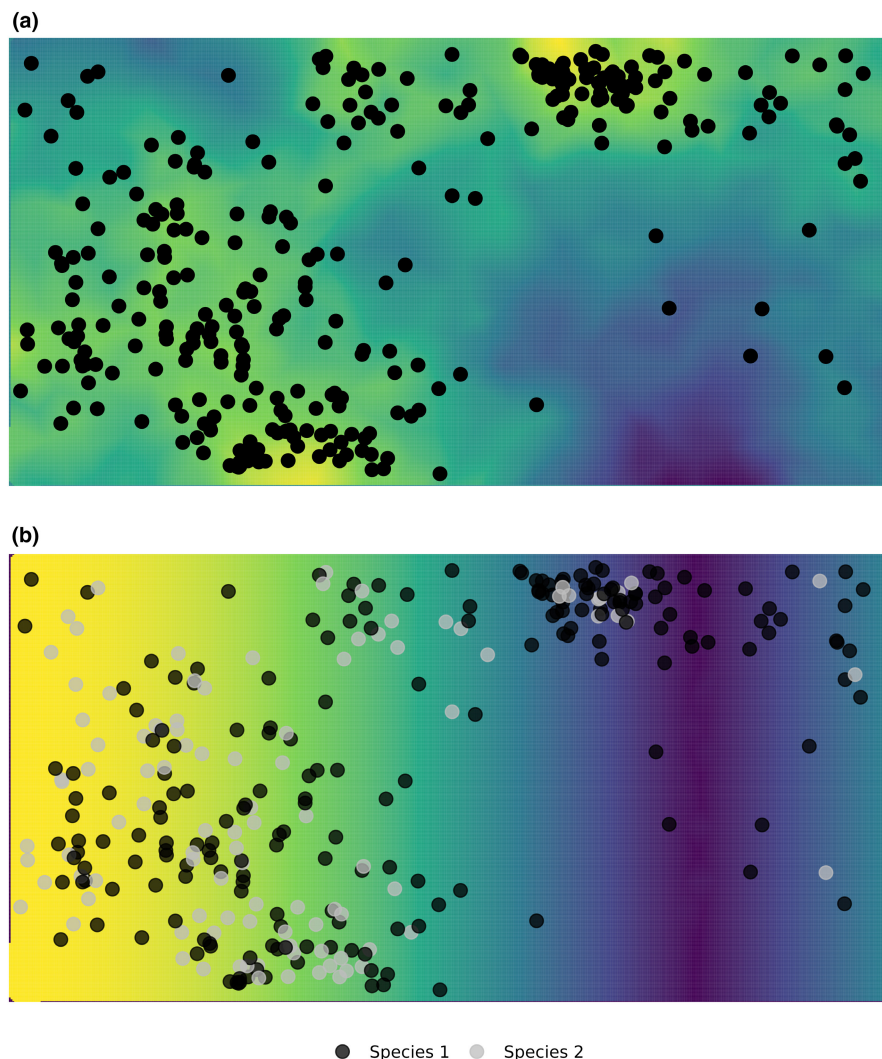
process (describing the number of points) with a (log) Gaussian random intensity process (describing the spatial dependence among points). The spatial distribution of points is conditionally independent given the Gaussian random intensity process, otherwise called a GRF.

Spatially clustered point pattern data occur often in ecological and environmental data. These clusters can be a result of many reasons (e.g., habitat preference, environmental covariates, over-sampling), which may be known and measurable or not. Often, an ecologist's interest lies in inferring the effect of environmental covariates while accounting (somehow) for the remaining unexplained spatial structure. Fitting a LGCP enables the effect of covariates to be inferred while the GRF accounts for any unexplained residual spatial dependence (e.g., clustering).

Figure 2 plot (a) shows a realisation of a LGCP overlain on the simulated Gaussian random intensity process. Here, larger values of the field are yellow, these higher values gives rise to more clustered points; whereas cooler areas (i.e., blue) give rise to more sparsely distributed points. In practice, it is only the points we observe and from their distribution we estimate the assumed latent GRF.

Mathematically, a LGCP is a specific type of Cox process (Cox & Isham, 1980) based on a stationary and isotropic Gaussian random field, $G(\mathbf{x})$, with mean μ , variance-covariance matrix $\mathbf{Q}^{-1}$

FIGURE 2 Plot (a) LGCP where the points are a realisation of an inhomogeneous Poisson process conditionally independent of the latent GRF. Parameters of the GRF control the expected spatial patterning of the points (i.e., the degree of clustering). Higher values of the GRF (here in yellow) gives rise to more clustered points, whereas cooler areas (here in blue) give rise to sparsely distributed points. Plot (b) A marked version of the LGCP shown in plot (a) where the chance of a Species 2 (grey) depends on the latent Gaussian intensity process in plot (a) as well as the surface in plot (b). That is, we are more likely to observe Species 2 where (1) there are more points, a process governed by the intensity surface in plot (a), and (2) the further to the left we are, this governed by the mark surface shown in plot (b).



and covariance function C_Z , such that the random intensity surface is given by

$$\Lambda(\mathbf{s}) = \exp(\mathbf{X}\beta + G(\mathbf{s})) \quad (3)$$

for design matrix $\mathbf{X}$ and coefficients β (Ripley, 1981). It is conventional to use the Matérn covariance function to define the covariance of the random field, which has two parameters ψ and κ . Typically, functions of ψ and κ are reported as the spatial range $r = \frac{\sqrt{8}}{\kappa}$ and standard deviation $\sigma = \frac{1}{\sqrt{4\pi\kappa^2\psi^2}}$. Under INLA methodology the *practical* range is defined as the distance such that the correlation is ~ 0.139 (Krainski et al., 2018). This range is given as $r\sqrt{8v} = \frac{\sqrt{8v}}{\kappa}$ where r is the spatial range (as above) and v is the smoothness parameter, which can be thought of as controlling the degree of change (spikiness) in the field.

As $\Lambda(\mathbf{x})$, above, is unobserved, the evaluation of the model likelihood involves integration over the infinite-dimensional distribution of $\Lambda(\mathbf{x})$, which typically results in a huge computational expense. An elegant solution is based on approximating the continuously indexed Gaussian field by a GMRF Rue and Tjelmeland (2002). More recently, this approach became particularly practical when Lindgren et al. (2011)

derived a solution based on a Stochastic Partial Differential Equation (SPDE) whose solution is a Gaussian field with Matérn correlation. They found that, for certain parameter values, it was possible to represent this Gaussian Field by a GMRF, by using the Finite Element Method to provide a solution to a SPDE. To make the GMRF approximation valid, v is fixed as $\alpha = v + \frac{d}{2}$ with $\alpha = 2$ (as Krainski et al. (2018)) in 2D ($d = 2$), therefore $v = 1$. We use this approach in *stelfi* to approximate the assumed Gaussian random fields. This requires the construction of a Delaunay triangulation for the spatial domain, which forms the basis of the approximation. For this, *stelfi* uses exported functionality from INLA (Lindgren & Rue, 2015).

A spatiotemporal LGCP model, typically, uses auto-regressive temporal dependence with an arbitrary number of time knots, $i = 1, \dots, n$ (e.g., month index):

$$\Lambda_i(\mathbf{s}) = \exp(\mathbf{X}\beta + G_i(\mathbf{s}) + \epsilon)$$

Here $\Lambda_i(\mathbf{s})$ is the field intensity at time knot i and $G_i(\mathbf{s})$ the GMRF at the same time knot. Each $G_i(\mathbf{s})$ shares common values for ψ and κ , and successive random fields are correlated following $G_i(\mathbf{s}) = \rho G_{i-1}(\mathbf{s}) + \epsilon_i$, where $\rho \in [-1, 1]$.

2.4 | A marked log-Gaussian Cox process

Each point in a spatial pattern may have one or more associated marks (characteristics). In that case, we are interested not only in the spatial intensity of the points but also in the spatial distribution of the marks (e.g., species) and the dependence between the marks and the spatial intensity. Here, we can again assume a log-Gaussian Cox model for the point pattern (as in the previous section), but also jointly model the marks. This is a useful model when the marks are not independent of clusters (e.g., certain species of shrub may be more likely to cluster together), and yet we are interested in the mark distribution independently of the inter-point dependence.

Figure 2 plot (b) shows a realisation of a marked LGCP where the colours indicate the species of each point, where points could be animals or plants etc. In Figure 2, the chance of a Species 2 (grey) depends on the latent Gaussian intensity process in plot (a) as well as the mark-specific surface in plot (b). This means that we are more likely to observe Species 2 where there are more points, a process governed by the GRF, and the further towards the left we are, this governed by the mark-specific surface.

A marked LGCP can be written as an extension to Equation (3), where

$$\Lambda(\mathbf{s}) = \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{G}(\mathbf{s}) + \epsilon),$$

$$M_j(\mathbf{s}) = f^{-1}((\mathbf{X}\boldsymbol{\beta})_{m_j} + G_{m_j}(\mathbf{s}) + \alpha_{m_j} \mathbf{G}(\mathbf{s}) + \epsilon_{m_j}). \quad (4)$$

Here, α_{m_j} are coefficient(s) linking the point process and the n_{mark} mark(s), $m_j (j = 1, \dots, n_{\text{mark}})$. The form of $M_j(\mathbf{s})$ and f^{-1} depends on the assumed distribution of the marks. This joint model can be likened to the preferential sampling model proposed by Diggle and Su (2010) where the log-Gaussian Cox model assumed for the point pattern is modelled jointly alongside the response. The response referred to in the preferential sampling model are, in the model above, the marks. The options available for the mark distribution in `stelfi` are given below.

1. If $m_j \sim \text{Normal}(\mu_j(\mathbf{s}), \sigma_j)$ then $M_j(\mathbf{s}) = \mu_j(\mathbf{s})$ and $f^{-1} = I$ (user must supply the standard deviation σ_j),
2. If $m_j \sim \text{Poisson}(\Lambda_j(\mathbf{s}))$ then $M_j(\mathbf{s}) = \Lambda_j(\mathbf{s})$ and $f^{-1} = \exp$,
3. If $m_j \sim \text{Binomial}(n_j, p_j(\mathbf{s}))$ then $M_j(\mathbf{s}) = p_j(\mathbf{s})$ and $f^{-1} = \text{logit}$ (user must supply the number of trials n_j), and
4. If $m_j \sim \text{Gamma}(\text{shape}_j(\mathbf{s}), \text{scale}_j)$ then $M_j(\mathbf{s}) = \text{shape}_j(\mathbf{s})$ and $f^{-1} = \log$ (user must supply the log of the scale parameter $\log(\text{scale}_j)$).

2.5 | A spatiotemporal self-exciting process

In this section we introduce a spatiotemporal self-exciting process, which combines some of the models outlined in the previous sections. Such a model would enable the degree of spatiotemporal self-excitement (e.g., that an event immediately increases of a future

event in nearby space) to be inferred in addition to the effect of spatial and temporal covariates.

The Hawkes process is traditionally formulated as a temporal point process; however, we can also extend this to a spatiotemporal version. The spatiotemporal Hawkes processes fitted by `stelfi` have the temporal self-excitation following an exponential decay function. The self-excitation over space follows a Gaussian distribution centred on the triggering event. There are two formulations of this model. The default is that the Gaussian function has a fixed covariance matrix, independent of time. Alternatively, the covariance can be directly proportional to time, meaning that the self-excitement radiates from the centre over time. This can be appropriate when the mechanism causing self-excitement travels over space; however, it is very memory-intensive. The spatiotemporal intensity function defined in `stelfi` is given by

$$\lambda(\mathbf{s}, t) = \mu + \mathbf{G}(\mathbf{s}) + \alpha \sum_{i: \tau_i < t} (\exp(-\beta^*(t - \tau_i))) K_i(\mathbf{s} - \mathbf{x}_i, t - \tau_i) + \epsilon. \quad (5)$$

where μ is the background rate, β is the rate of temporal decay, α is the increase in intensity after an event, τ_i are the event times, and $\mathbf{x}_i$ are the event locations (in 2D Euclidean space). The error term is given by ϵ . As in Equation (3) $\mathbf{G}(\mathbf{s})$ is a Gaussian random field, the inclusion of which is optional. The spatial self-excitement kernel is given by $K_i(\mathbf{s} - \mathbf{x}_i, t - \tau_i) \sim \text{Normal}(0, \mathbf{Q}^{-1})$; this can either be time-

independent where $\mathbf{Q}^{-1} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$ or time-dependent

where $\mathbf{Q}^{-1} = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} \times (t_j - t_i)$ for $t_j > t_i$.

2.6 | Example

The BFRO is an organisation dedicated to investigating mystery of Sasquatch (bigfoot) (BFRO, 1995). The organisation collects reports from the general public detailing (supposed) Sasquatch sightings. Alongside a brief description of the incident each sighting occurrence is geocoded (latitude and longitude) and has a timestamp, with resolution varying depending on report. Although the existence of Sasquatch is unlikely, the structure of the data is akin to many other citizen science datasets. In addition, the questions and issues that arise from such data are equivalent to those ecologists have in dealing with sightings data for non-mythical creatures. For example, sightings often cluster in time and space due to a variety of reasons. These reasons may be environmental, ecological or simply be an artefact of the observation process (e.g., people are more likely to visit areas where a certain species has previously been sighted and are therefore more likely to observe it). The interest is often to tease apart or account for the multiple reasons behind the clusters (or otherwise) of sightings. In this section, we illustrate the fitting of the most common models outlined in the sections above using data collated by Renner (2021) obtained from the BFRO containing the temporal

and spatial locations of claimed sightings of Sasquatch across the contiguous USA from 2000 to 2006. All the required R code for the analysis summarised in this section as well as more detail about the Sasquatch data is given in Appendix S1 as well as the package website (<https://cmjt.github.io/stelfi/>). In Appendix 1, we show via simulation the performance of the model fitting functions described below.

2.7 | Self-exciting sightings using `fit_hawkes()`

From 2000 up until 2006 there were $n = 972$ recorded Sasquatch sightings over $T = 2188$ days in the contiguous USA, see plot (a) in Figure 3. These sightings data were recorded at a daily resolution; however, the Hawkes process assumes unique occurrence times. Therefore, to fit the model we jitter each observation slightly (by no more than 24h). As the data cover such a long span this makes little difference in practice; however, in general, great care should be taken when including a jitter.

Hawkes processes are particularly useful tools in inferring the, potentially, contagious nature of clustered event-type data. For example, we may be interested in inferring to what extent the claimed sighting of bigfoot incites another in the near future. To fit a Hawkes process given by Equation (1) using `stelfi` to the Sasquatch data we call

```
fit_hawkes(times, parameters)
```

where `times` is a vector of Sasquatch sighting times (i.e., temporal locations as described above) and `parameters` is a vector of parameter starting values. An additional argument, `marks` is not supplied, which means by default $m(t)$ in Equation (1) is fixed at 1. The fitted model, estimated self-exciting temporal intensity of sightings, can be seen in plot (a) Figure 3, where the rug plot on the x-axis shows the temporal location of recorded sightings.

Table 1 compares the parameter estimates and standard errors returned by `stelfi` with `emhawkes` and `hawkesbow` (see the Appendix S1 for the required R code). From these estimated values, we estimate the expected background rate of sightings (i.e., independent sightings) as $\hat{\mu}T = 0.12 \times 2188 \sim 263$, which indicates that ~ 263 sightings were principal (independent) sightings and that the remaining were due to self-excitement. The expected number of sightings “triggered” by anyone sighting is estimated as $\hat{\alpha} = \frac{0.06}{0.09} = \frac{2}{3}$ and the expected number of descendants per sighting is estimated as $\frac{\beta}{\beta - \alpha} = \frac{0.09}{0.09 - 0.06} = 3$. The rate of decay for the self-excitement is estimated as $\frac{1}{\hat{\beta}} = \frac{1}{0.09} \sim 11$ days, indicating that after ~ 11 days a sighting is likely independent of any historic sighting.

Giving the fitted model object to the `stelfi` functions `show_hawkes()` and `show_hawkes_GOF()` will plot the fitted model and a range of goodness-of-fit plots, respectively. In addition, if `return_values=TRUE` is set, the transformed inter-arrival times from the fitted models are returned, see the package website (<https://cmjt.github.io/stelfi/>) for examples.

2.8 | Spatial distribution of Sasquatch sightings using `fit_lgcp()`

The $n = 972$ Sasquatch sightings occurred within an area ($A \sim 7,771,155\text{km}^2$) (the contiguous USA, plot (a) in Figure 3). This shows that sightings are, spatially, quite sparse at ~ 0.000125 per km^2 over the 5 years. Most sightings occur in Washington and north-west Oregon (to be expected as Bigfoot is purported to inhabit the wild and forested areas of Oregon and the West Coast of North America), with another dense spot near the Texas–Louisiana border.

To infer the spatial distribution of bigfoot (sightings) we fit a LGCP given by Equation (3) with $\mathbf{X} = \mathbf{1}$ and $\beta = \beta_0$ (i.e., an intercept-only fixed effect) in `stelfi` via

```
fit_lgcp(locs, sf, smesh, parameters)
```

where `locs` is a data frame of 2D locations, `sf` a simple features polygon of the domain/window, `smesh` a Delaunay triangulation covering the domain, and `parameters` a vector of parameter starting values.

We compare the parameter estimates returned by `stelfi` to `inlabru` in Table 1 (see the Appendix S1 for the required R code). Both packages use the SPDE approach to estimate the latent field (Lindgren et al., 2011); however, `stelfi` uses a frequentist framework with custom-written TMB templates and `inlabru` uses a Bayesian framework.

To estimate the total number of events, integration weights are required. These weights are the area surrounding the mesh nodes; the expected number of events at each node can be thought of as being proportional to the weights (Simpson et al., 2016). The weights used are the areas of the Voronoi cell centred on each mesh node (called the *dual mesh* in the INLA framework). The inbuilt `get_weights()` can be used in `stelfi` to extract the weights used given a model object and a Delaunay triangulation. To plot these weights set `plot=TRUE`, this results in the top figure of plot (c) in Figure 3, which shows the Delaunay triangulation (white) and the resulting Voronoi tessellation (grey). Each Voronoi cell is coloured according to the integration weight used in the model; note that all cells outwith the contiguous USA land mass are set to zero and therefore do not contribute to the model. Using these integration weights and the estimated parameters (Table 1) we estimate the total number of events using $\sum_{k=1}^m w_k \exp(\beta_0 + g_k)$ where m are the number of mesh nodes, w_k is the weight and g_k the estimated mean of the random field at the k th mesh node; from this, we get an estimate of 970 events using `stelfi` and 987 using `inlabru` (observed $n = 972$ sightings).

Figure 3 plot (c) shows the estimated intensity surface obtained by each package. To return the estimated values of the random field at each mesh node from `stelfi` we use `get_fields()`, providing the fitted model object, `fit`, and the mesh, `smesh`. There are additional options to return the estimated standard deviation of the field(s) with `sd=TRUE` and to plot the returned values with `plot=TRUE`. The estimated intensity surface (middle plot) is plotted using `show_lambda` with the additional arguments `sf` and `clip=TRUE`, to plot only the values within the domain. The plots

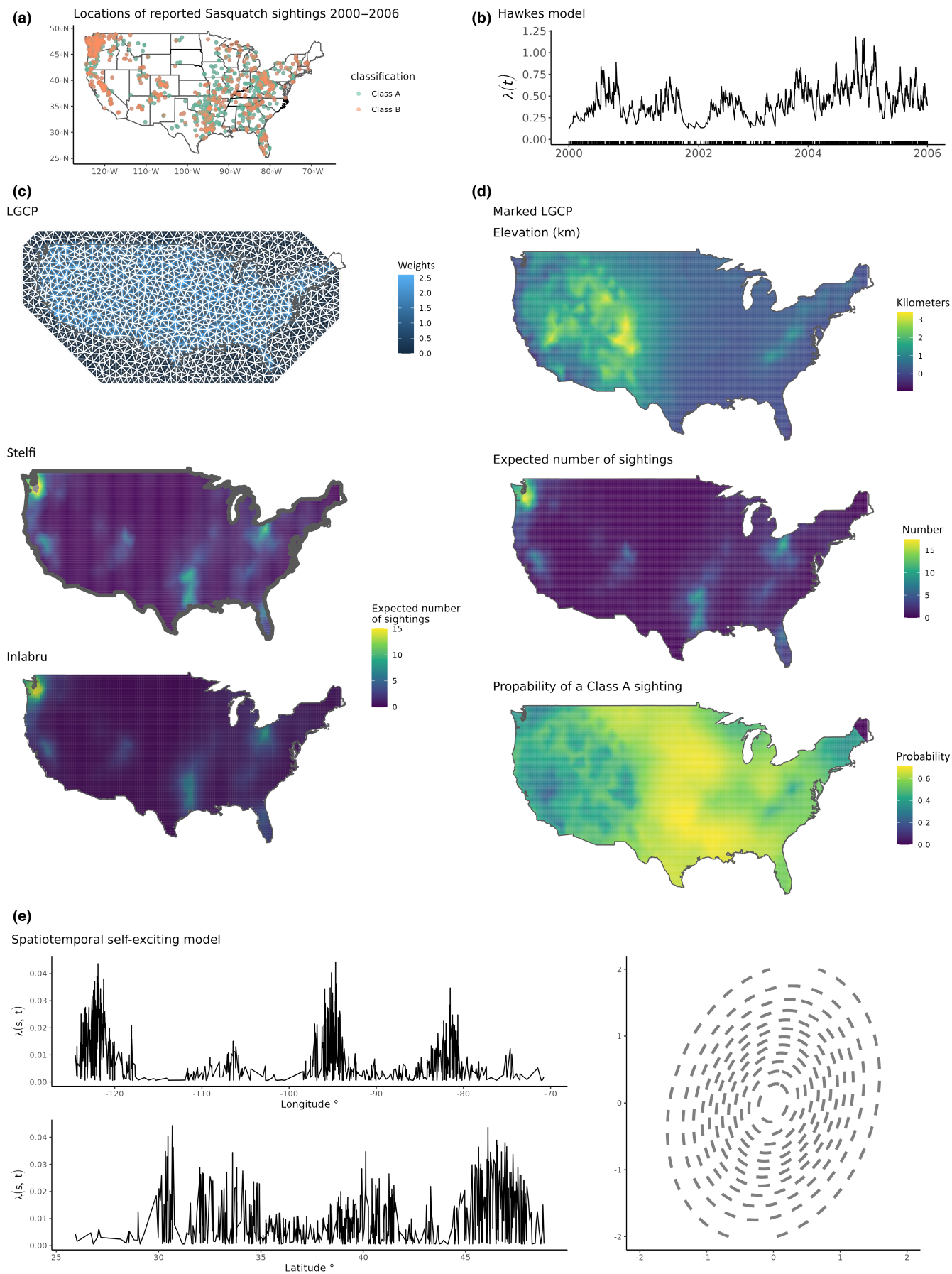


FIGURE 3 Plot (a) shows the observed point pattern, points are coloured by marks. Plot (b) shows the fitted Hawkes intensity (black line) given by Equation (1) to the temporal point pattern (rug plot) of Sasquatch sighting times. Plot (c) shows the Delaunay triangulation (white) and the resulting Voronoi tessellation (grey) where each Voronoi cell is coloured according to the integration weight used in the model; In addition, the estimated intensity surface obtained by `stelfi` and for comparison `inlabru` are shown. Plot (d) shows the components of the marked LGCP given in Equation (6): the elevation covariate, the fitted intensity surface, and the estimated mark process. Plot (e) shows the spatiotemporal intensity, over Latitude and Longitude, given by Equation (5) and the estimated Gaussian kernel density for Q^{-1} shown in Table 1.

are `ggplot2` objects and therefore additional `geom`'s can be added as layers. See the Appendix S1 for the required `inlabru` code for the comparisons given.

2.9 | A marked LGCP using `fit_mlgcp()`

Categorical marks of ecological objects are commonplace (e.g., species, sex, classification) and often there is interest in inferring the point-mark dependence and the mark-specific distribution (i.e., independent of point location), see Figure 2. The Bigfoot Field Researchers Organisation (BFRO) (BFRO, 1995) categorise Sasquatch sightings into three classes: (1) Class A, clear; (2) Class B, observed at a great distance or in poor lighting conditions; and Class C, second- and third-hand reports, or stories with an untraceable source.

To illustrate a marked LGCP we consider only Class A and Class B sightings ($n = 971$) and jointly model Sasquatch sightings alongside the probability of that sighting being of Class A (clear). Doing so will allow us to infer the, potentially, unexplained process driving clear sightings (e.g., lack of foliage). To do this, we set $m_i = 1$ if the sighting is Class A and $m_i = 0$ if the sighting is Class B for $i = 1, \dots, 971$. The joint model we fit is a special case of Equation (4):

$$\begin{aligned} \Lambda(s) &= \exp(\beta_0 + \beta_1 x_{\text{elev}}(s) + G(s) + \epsilon) \\ \text{logit}(p(s))^{-1} &= \beta_{0m} + \beta_{1m}^m x_{\text{elev}}(s) + G_m(s) + \alpha_m G(s) + \epsilon_m \end{aligned} \quad (6)$$

where $m(s) \sim \text{Bernoulli}(p(s))$ and the spatial intensity of all sightings is, as previously, $\Lambda(s)$. A spatial covariate, $x_{\text{elev}}(s)$ the elevation in kilometres obtained from the R package `elevatr` (Hollister et al., 2021), is included in each linear predictor (see plot (d) in Figure 3). Different coefficients are estimated for the LGCP and the mark process (β_1 and β_{1m}^m respectively). It is a reasonable supposition that elevation might affect the clarity of the sighting as well as their number and therefore it is included, separately, as a covariate in both linear predictors. The shared GMRF $G(s)$ represents the spatial autocorrelation of the sighting locations, which may also contribute to the mark process according to the estimated parameter α_m . The GMRF $G_m(x)$ is unique to the mark and may be thought of as representing any spatial autocorrelation not explained by the point locations; this interpretation is somewhat ad hoc however and care should be taken.

To fit this model in `stelfi`, we use

```
fit_mlgcp(locs, marks, sf, smesh, parameters,
  methods = 2, fields = 1, covariates = covariates,
  pp_covariates = 1, marks_covariates = 1)
```

where, in addition to the arguments we supply to `fit_mlgcp()` we set `methods=2` (Binomial distribution assumed for the marks with the number of trials = 1, Bernoulli), and `fields=1` (to include mark-specific random field). We also supply a vector `marks`, the mark values at each location and `covariates` the elevation in kilometres at each mesh node obtained from the R package `elevatr` (Hollister et al., 2021). Setting `pp_covariates=1` and `marks_covariates=1` means that the first column of `covariates` will be included in the linear predictor for both the LGCP and the mark.

Table 1 gives the estimated parameter values. The magnitude of the contribution of the shared random field to the mark is given by the parameter α_m which is estimated as 0.044 (standard error 0.063). This indicates that there is no point-mark dependence because the estimated standard error is greater than the distance between the estimated coefficient value and zero. That is, the probability of a Class A sighting does not seem to change in relation to the spatial intensity of sightings. The mark-specific elevation coefficient is estimated as -0.372 (standard error 0.178) indicating that the probability of a Class A sighting (slightly) decreases as elevation increases. This seems reasonable, people are perhaps less likely to get a clear sighting the greater the elevation due to rougher terrain etc. Looking at the estimated mark process in Figure 3 plot (d) we see that the probability of a Class A sighting is greatest in the eastern states. The elevation coefficient for the LGCP is estimated as 0.113 (standard error 0.251) indicating that sightings increase as elevation increases. Again, this seems reasonable, sightings are more prolific the greater the elevation/the more remote the terrain, see Figure 3 plot (d).

2.10 | Self-exciting spatiotemporal sightings using `fit_stelfi()`

Fitting a spatiotemporal self-exciting model requires that the times of the events be unique, therefore, we have removed duplicated daily sightings for the Sasquatch data. Not something that is advised in practice. However, in the interest of demonstration, we do so here resulting in 677 sightings on unique days; see the Appendix S1 for the required R code.

To fit a time-independent model as Equation (5) we use the `stelfi` function

```
fit_stelfi(times, locs, sf, smesh, parameters)
```

where the arguments are as above. In addition to these typical arguments, there are two logical arguments to `fit_stelfi` that control

TABLE 1 Table showing the estimated parameter values and standard errors for the fitted models discussed.

Model	Parameter(s)	stelfi		emhawkes		hawkesbow		inlabru	
		Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Hawkes	μ	0.12	0.04	0.12	0.05	0.12	0.04	–	–
	α	0.06	0.03	0.07	0.03	–	–	–	–
	$\frac{\alpha}{\beta}$	–	–	–	–	0.73	0.10	–	–
	β	0.09	0.05	0.09	0.06	0.09	0.05	–	–
	log-likelihood	–1681.156		–1676.068		–1681.156		–	
LGCP	β_0	–0.77	0.36	–	–	–	–	–0.49	0.25
	r	7.34	1.11	–	–	–	–	7.94	0.83
	σ	1.35	0.15	–	–	–	–	0.87	0.06
	system. time() \$elapsed	25.37		–		–		114.83	
Marked LGCP	β_0	–0.823	0.358	–	–	–	–	–	–
	β_1	0.114	0.251	–	–	–	–	–	–
	β_0^m	0.227	0.284	–	–	–	–	–	–
	β_1^m	–0.372	0.178	–	–	–	–	–	–
Self-exciting spatiotemporal model	α_m	0.044	0.063	–	–	–	–	–	–
	μ	0.0006	0.0001	–	–	–	–	–	–
	α	0.0215	0.0028	–	–	–	–	–	–
	β	0.0215	0.0028	–	–	–	–	–	–
	σ_x	0.7060	0.0518	–	–	–	–	–	–
	σ_y	0.9132	0.0913	–	–	–	–	–	–
	ρ	0.1999	0.1226	–	–	–	–	–	–
	Q^{-1}	$\begin{bmatrix} 0.498 & 0.129 \\ 0.129 & 0.834 \end{bmatrix}$		–		–		–	

Note: Hawkes: Estimated parameter and log-likelihood values of a Hawkes process, given by Equation (1), modelling the reported temporal occurrence of Sasquatch sightings. These are shown for the packages `stelfi`, `emhawkes` and `hawkesbow`. Note that `hawkesbow` returns the estimate of $\frac{\alpha}{\beta}$ and the standard errors shown are calculated from the returned Hessian matrix H (i.e., $\sqrt{\text{diag}((-H)^{-1})}$). LGCP: Estimated intercept, β_0 , and GMRF parameter values (range r and standard deviation σ) for the LGCP process (Equation 3) fitted to reported Sasquatch sighting locations. Estimates and standard errors are shown for both `stelfi` and `inlabru` along with the time elapsed to fit the models. Marked LGCP: Estimated parameter of the joint model given by Equation (6) fitted using `stelfi::fit_mlhcp()`. Self-exciting spatiotemporal model: Estimated parameter values, standard errors, and the spatiotemporal precision matrix for the model given by Equation (5) fitted using `stelfi::fit_stelfi()`.

the form of spatiotemporal self-excitement. Setting `time_independent=TRUE` (default `FALSE`) makes the self-excitement kernel time-dependent, see the package website for more details (<https://cmjt.github.io/stelfi/>). Setting `GMRF=TRUE` (default `FALSE`) adds a Gaussian Markov random field onto the baseline spatiotemporal intensity.

Parameter estimates for the fitted model are given in Table 1. Figure 3 plot (e) show the spatiotemporal intensity, over Latitude and Longitude, as well as the bivariate Gaussian density with Q^{-1} as shown in Table 1. From this, we can see that the estimated spatiotemporal intensity over longitude shows spikes of sightings clustered around 120° W, 95° W and 80° W, correspondingly over latitude the intensity is relatively similar, however, 30° N and 45° N seem to have a greater intensity. These locations correspond to Washington, Texas and Ohio, matching the observed data structure over those states. The contour plot shows the estimated Gaussian

spatial kernel in Equation (5); from the elliptical shape this shows that the spatial dependency is more dispersed across latitude (y-axis) than longitude (x-axis).

3 | CONCLUSIONS

The R package `stelfi` offers the functionality to fit a variety of Hawkes and log-Gaussian Cox process models via TMB (Kristensen et al., 2016). The package is available on CRAN, with the development version on GitHub (<https://github.com/cmjt/stelfi>). The package version (v1.0.1) used in this manuscript has the persistent identifier <https://doi.org/10.5281/zenodo.10515700>.

Modelling point pattern data is a huge part of ecological and environmental statistics. The novelty of the package is twofold (1) custom-written C++ templates take advantage of TMB, fitting

models via maximum likelihood; (2) it facilitates the fitting of temporal and spatiotemporal self-exciting point process models.

We demonstrate the core functionality of `stelfi` using Sasquatch sightings data. These data are used for illustration purposes only, yet the structure is typical of data often seen in ecology and environmental science. Where possible, we demonstrate the performance of `stelfi` compared to the R packages that fit similar models via different means.

This package offers a much-needed addition to the point process field. It provides an alternative to existing model fitting frameworks and includes functionality that extends the currently available methods. We foresee the package becoming requisite in the fields of ecological and environmental modelling and beyond.

AUTHOR CONTRIBUTIONS

Charlotte M. Jones-Todd: Conceptualization (lead); formal analysis (lead); funding acquisition (lead); methodology (lead); project administration (lead); software (lead); writing – original draft (lead). **Alec B. M. van Helsdingen:** Software (supporting); writing – review and editing (supporting).

ACKNOWLEDGEMENTS

The authors wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high-performance computing facilities, consulting support and/or training services as part of this research. New Zealand's national facilities are provided by NeSI and are jointly funded by the NeSI collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme, URL <https://www.nesi.org.nz>. Open access publishing facilitated by The University of Auckland, as part of the Wiley - The University of Auckland agreement via the Council of Australian University Librarians.

FUNDING INFORMATION

This work was supported by Marsden Fund proposal UOA 3723517 and Asian Office of Aerospace Research & Development grant FA2386-21-1-4028.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data used in this manuscript are shipped with the R package `stelfi` available from the Comprehensive R Archive Network (CRAN) (<https://cran.r-project.org/package=stelfi>). Furthermore, the R package code is open source and available from the GitHub repository <https://github.com/cmjt/stelfi> (<https://doi.org/10.5281/zenodo.10515700>). All model fitting code and package version information etc. is given in Appendix S1, <https://doi.org/10.5281/zenodo.10516195>.

ORCID

Charlotte M. Jones-Todd  <https://orcid.org/0000-0003-1201-2781>

REFERENCES

- Bachl, F. E., Lindgren, F., Borchers, D. L., & Illian, J. B. (2019). inlabru: An R package for Bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution*, 10(6), 760–766. <https://doi.org/10.1111/2041-210X.13168>
- Bacry, E., Mastromatteo, I., & Muzy, J.-F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1), 1550005. <https://doi.org/10.1142/S2382626615500057>
- Baddeley, A., Rubak, E., & Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. Chapman and Hall/CRC Press.
- Ben-Said, M. (2021). Spatial point-pattern analysis as a powerful tool in identifying pattern-process relationships in plant ecology: An updated review. *Ecological Process*, 10(56), 1–23. <https://doi.org/10.1186/s13717-021-00314-4>
- BFRO. (1995). *The Bigfoot Field Researchers Organization (BFRO)*. <https://www.bfro.net/>
- Cheysson, F. (2021). *hawkesbow: Estimation of Hawkes processes from binned observations*. R package version 1.0.2. <https://CRAN.R-project.org/package=hawkesbow>
- Cox, D. R., & Isham, V. (1980). *Point processes*. Chapman and Hall/CRC Press.
- Diggle, P. J. (2013). *Statistical analysis of spatial and spatio-temporal point patterns*. Chapman and Hall/CRC Press.
- Diggle, P. J., R. M., & Su, T.-L. (2010). Geostatistical inference under preferential sampling. *Journal of the Royal Statistical Society: Series C: Applied Statistics*, 59, 191–232. <https://doi.org/10.1111/j.1467-9876.2009.00701.x>
- Gupta, A., Farajtabar, M., Dilkina, B., & Zha, H. (2018). Discrete interventions in Hawkes processes with applications in invasive species management. In *Proceedings of the twenty-seventh international joint conference on artificial intelligence, IJCAI-18* (pp. 3385–3392). International Joint Conferences on artificial Intelligence Organization. <https://doi.org/10.24963/ijcai.2018/470>
- Hawkes, A. (1971a). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1), 83–90. <https://doi.org/10.2307/2334319>
- Hawkes, A. G. (1971b). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, 33(3), 438–443.
- Hawkes, A. G. (2018). Hawkes processes and their applications to finance: A review. *Quantitative Finance*, 18(2), 193–198. <https://doi.org/10.1080/14697688.2017.1403131>
- Hollister, J., Shah, T., Robitaille, A. L., Beck, M. W., & Johnson, M. (2021). *elevatr: Access elevation data from various APIs*. R package version 0.4.2. <https://CRAN.R-project.org/package=elevatr>
- Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns*. John Wiley & Sons.
- Illian, J. B., Sørbye, S. H., & Rue, H. (2012). A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *The Annals of Applied Statistics*, 6(4), 1499–1530.
- Jones-Todd, C. M., & van Helsdingen, A. (2023). *stelfi: Hawkes and log-Gaussian Cox point processes using template model builder*. R package version 1.0.0. <https://CRAN.R-project.org/package=stelfi>
- Krainski, E., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F., & Rue, H. (2018). *Advanced spatial modelling with stochastic partial differential equations using R and INLA*. Chapman and Hall/CRC Press.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software*, 70(5), 1–21. <https://doi.org/10.18637/jss.v070.i05>
- Lee, K. (2021). *emhawkes: Exponential multivariate Hawkes model*. R package version 0.9.5. <https://CRAN.R-project.org/package=emhawkes>
- Lindgren, F., & Rue, H. (2015). Bayesian spatial modelling with R-INLA. *Journal of Statistical Software*, 63(19), 1–25. <https://doi.org/10.18637/jss.v063.i19>

- Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between gaussian fields and gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, 73(4), 423–498. <https://doi.org/10.1111/j.1467-9868.2011.00777.x>
- Møller, J., Syversveen, A. R., & Waagepetersen, R. P. (1998). Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25(3), 451–482.
- Nakagawa, T., Subbey, S., & Solvang, H. (2019). Integrating Hawkes process and biomass models to capture impulsive population dynamics. *Dynamics of Continuous, Discrete and Impulsive Systems Series B (Applications and Algorithms)*, 26(3), 153–170.
- Naylor, M., & Serafini, F. (2023). ETAS.inlabru: Temporal Hawkes. v1.0.1. <https://doi.org/10.5281/zenodo.7515785>
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 3(401), 9–27. <https://doi.org/10.2307/2288914>
- Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1), 145–155. <https://doi.org/10.1007/BF02480272>
- Park, J., Schoenberg, F. P., Bertozzi, A. L., & Brantingham, P. J. (2021). Investigating clustering and violence interruption in gang-related violent crime data using spatial-temporal point processes with covariates. *Journal of the American Statistical Association*, 116(536), 1674–1687. <https://doi.org/10.1080/01621459.2021.1898408>
- Renner, T. (2021). Data from: Bigfoot sightings. *Data World*. <https://data.world/timothyrenner/bfro-sightings-data>
- Ripley, B. D. (1981). *Spatial statistics*. Wiley.
- Ripley, B. D. (1987). Spatial point pattern analysis in ecology. In P. Legendre & L. Legendre (Eds.), *Developments in numerical ecology* (pp. 407–429). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-70880-0_11
- Robert, C. P., Casella, G., & Casella, G. (1999). *Monte Carlo statistical methods* (Vol. 2). Springer. <https://doi.org/10.1007/978-1-4757-4145-2>
- Ross, G. J. (2017). bayesianETAS: Bayesian estimation of the ETAS model for earthquake occurrences. R package version 1.0.3. <https://CRAN.R-project.org/package=bayesianETAS>
- Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, 71(2), 319–392. <https://doi.org/10.1111/j.1467-9868.2008.00700.x>
- Rue, H., & Tjelmeland, H. (2002). Fitting Gaussian Markov random fields to Gaussian fields. *Scandinavian Journal of Statistics*, 29(1), 31–49.
- Serafini, F., Lindgren, F., & Naylor, M. (2023). Approximation of Bayesian Hawkes process with inlabru. *Environmetrics*, 34(5), e2798. <https://doi.org/10.1002/env.2798>
- Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P., Díaz-Ávalos, C., & Rue, H. (2014). Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: The case of Catalonia, 1994–2008. *Environmental and Ecological Statistics*, 21(3), 531–563. <https://doi.org/10.1007/s10651-013-0267-y>
- Simpson, D., Illian, J. B., Lindgren, F., Sørbye, S. H., & Rue, H. (2016). Going off grid: Computationally efficient inference for log-Gaussian Cox processes. *Biometrika*, 103(1), 49–70.
- Skaug, H. J., & Fournier, D. A. (2006). Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. *Computational Statistics & Data Analysis*, 51(2), 699–709. <https://doi.org/10.1016/j.csda.2006.03.005>
- Soriano-Redondo, A., Jones-Todd, C. M., Bearhop, S., Hilton, G. M., Lock, L., Stanbury, A., Votier, S. C., & Illian, J. B. (2019). Understanding species distribution in dynamic populations: A new approach using spatio-temporal point process models. *Ecography*, 42(6), 1092–1102. <https://doi.org/10.1111/ecog.03771>
- Taylor, B. M., Davies, T. M., Rowlingson, B. S., & Diggle, P. J. (2013). lgc: An R package for inference with spatial and spatio-temporal log-Gaussian Cox processes. *Journal of Statistical Software*, 52, 1–40. <https://doi.org/10.18637/jss.v052.i04>
- Velázquez, E., Martínez, I., Getzin, S., Moloney, K. A., & Wiegand, T. (2016). An evaluation of the state of spatial point pattern analysis in ecology. *Ecography*, 39(11), 1042–1055. <https://doi.org/10.1111/ecog.01579>
- Wiegand, T., & Moloney, K. (2014). *A handbook of spatial point pattern analysis in ecology* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/b16195>
- Zaatour, R. (2014). hawkes: Hawkes process simulation and calibration toolkit. R package version 0.0.4. <https://CRAN.R-project.org/package=hawkes>
- Zhuang, J., & Mateu, J. (2019). A semiparametric spatiotemporal Hawkes-type point process model with periodic background for crime data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(3), 919–942. <https://doi.org/10.1111/rssa.12429>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Jones-Todd, C. M., & van Helsdingen, A. B. M. (2024). *stelfi*: An R package for fitting Hawkes and log-Gaussian Cox point process models. *Ecology and Evolution*, 14, e11005. <https://doi.org/10.1002/ece3.11005>

APPENDIX 1

SIMULATION STUDY

In this section, we show via simulation the performance of the model fitting functions described in the manuscript. In each case, we simulate 500 realisations of a temporal/spatial/spatiotemporal point process using, where applicable, the same parameter values estimated from fitting the model to the Sasquatch data. We use these parameter values so that the simulated data are a similar structure to the Sasquatch data modelled. We then fit the same (i.e., known true) model to each realisation and assess model performance using the estimated parameters from the fitted models.

Plot (a) in Figure A1 shows a single realisation of a temporal Hawkes process (Equation 1) with $\mu = 0.12$, $\alpha = 0.06$, and $\beta = 0.09$ over a period of $T = 2500$ days (similar to the time frame the Sasquatch data covers). Plot (b) in Figure A1 shows a single realisation of a simulated temporal bivariate Hawkes process (Equation (2)

where $N = 2$) with $\mu = (0.2, 0.2)$, $\alpha = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{pmatrix}$, and $\beta = (0.7, 0.7)$.

Here event type 1 (Stream 1) is shown in black and event type 2 (Stream 2) is shown in grey. Figure A2 shows the raw bias (i.e., for any parameter value θ , raw bias = $\hat{\theta} - \theta$) for the parameters of (a) a temporal Hawkes process and (b) a bivariate temporal Hawkes process. A total of 500 realisations are simulated and then the same model fitted to each realisation; Table A1 gives the parameter values

FIGURE A1 (a) Single realisation of a simulated temporal Hawkes process (Equation 1) with $\mu = 0.12$, $\alpha = 0.06$ and $\beta = 0.09$. The realisation is simulated over $T = 2500$ days, similar to the time frame the Sasquatch data covers. The corresponding Hawkes intensity, $\lambda(t)$, is shown by the solid line and each event is denoted by $\cdot$. (b) A single realisation of a simulated temporal bivariate Hawkes process (Equation 2) where $N = 2$ with $\mu = (0.2, 0.2)$, $\alpha = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{pmatrix}$, and $\beta = (0.7, 0.7)$. The Hawkes intensities are shown by the solid lines and each event is denoted by $\cdot$. The colours of the lines and the events indicate the two different streams.

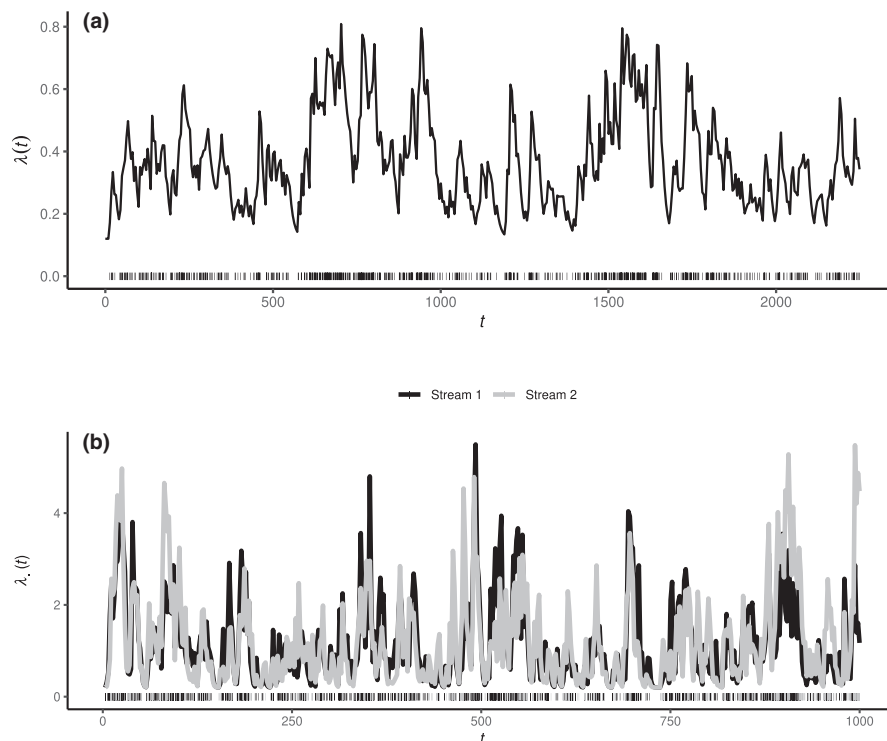
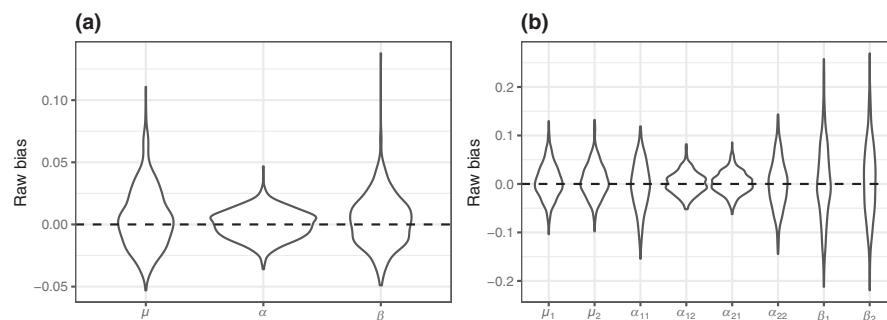


FIGURE A2 Raw bias (i.e., for any parameter value θ , raw bias = $\hat{\theta} - \theta$) for each simulation study (a) a temporal Hawkes process and (b) a bivariate temporal Hawkes process. Table A1 gives the parameter values used in the simulations as well as the percentage bias in each case.



used in the simulation as well as the percentage bias $\frac{\bar{\hat{\theta}} - \theta}{\theta}$, where $\bar{\hat{\theta}}$ is the average estimated parameter value across all fitted models) for each simulation.

Plot (a) in Figure A3 shows a single realisation of a simulated spatial GMRF (see Equation 3) over the contiguous USA, with parameter values $\beta = -0.76$, $\log(\tau) = -0.60$ and $\log(\kappa) = -0.95$ (i.e., $r = 7.34$ and $\sigma = 1.34$, values estimated in the corresponding Sasquatch model). During the simulation study 500 realisations of a LGCP (Equation 3) were simulated with the parameter values given in Table A1 and the same model (i.e., known true) fitted to each realisation. Plot (b) in Figure A3 shows the raw bias for each parameter.

Plot (a) in Figure A4 shows a single realisation of a simulated spatiotemporal Hawkes process (Equation 5) where the timestamp of each point is denoted by the size of the plotting character (the larger the later the timestamp). An arbitrary point in the pattern is denoted

X and the bivariate Gaussian kernel corresponding to the spatial self-excitation is shown by the coloured contours (i.e., a greater chance of a point in the near future in the lighter-coloured areas than the darker ones). The Gaussian kernel has parameter values $\sigma_x = 0.7$, $\sigma_y = 0.9$ and $\rho = 0.2$ corresponding to the estimated values from the fitted Sasquatch model. As previously, 500 realisations of the process were simulated with the parameter values given in Table A1 and the same model fitted to each realisation. Plot (b) in Figure A4 shows the raw bias for each parameter with the % bias shown in Table A1.

From plot (b) in Figure A4 and Table A1 we can see clear negative bias in the estimation of the self-excitement parameters. This is likely due to the weak spatial self-excitement process (i.e., relatively small values of α , σ_x , σ_y and ρ) inherent in the simulated data. This is evident in the coloured contours of plot (a) in Figure A4, which are relatively smooth and decay slowly over space.

Model	Parameter	Value	% Bias
Hawkes	μ	0.12	4.42
	α	0.06	0.18
	β	0.09	3.60
Multi. Hawkes	μ	(0.2,0.2)	(2.78,3.55)
	α	$\begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{bmatrix}$	$\begin{bmatrix} -0.73 & 1.63 \\ 2.59 & 0.03 \end{bmatrix}$
	β	(0.7,0.7)	(0.68,1.08)
LGCP	β	-0.76	-0.78
	$\log(\tau)$	-0.60	0.87
	$\log(\kappa)$	-0.95	-6.05
Self-exciting spatiotemporal model	μ	2	10
	$\frac{\alpha}{\beta}$	1	-46
	σ_x	0.7	-49
	σ_y	0.9	-53
	ρ	0.2	-35

TABLE A1 True parameter values and the calculated percentage bias ($\hat{\theta} - \theta / \theta$, where $\hat{\theta}$ is the average estimated parameter value across all fitted models) for each simulation study.

Note: In each case, a total of 500 realisations of the corresponding point process were simulated and then the same (i.e., known true) model fitted to each realisation.

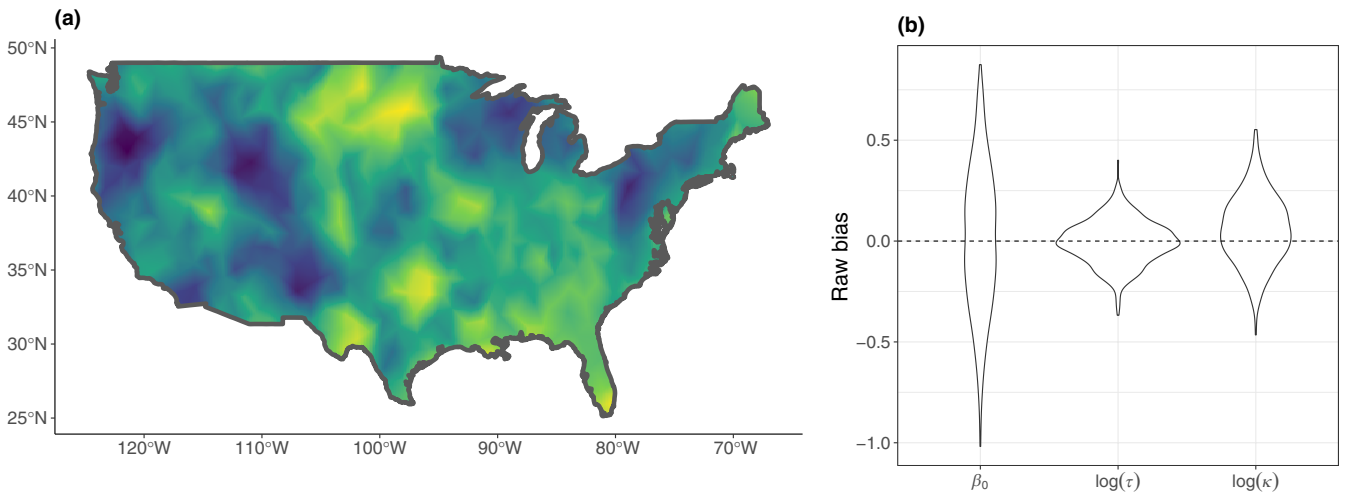


FIGURE A3 (a) Single realisation of a simulated spatial GMRF over the contiguous USA, with parameter values $\beta = -0.76$, $\log(\tau) = -0.60$ and $\log(\kappa) = -0.95$ (i.e., $r = 7.34$ and $\sigma = 1.34$, values estimated in the corresponding Sasquatch model). (b) The raw bias (i.e., for any parameter value θ , raw bias = $\hat{\theta} - \theta$) for each parameter of the LGCP.

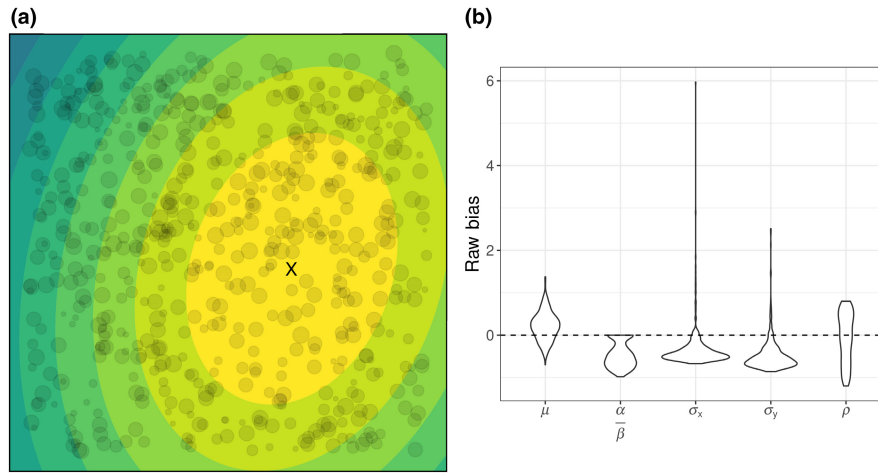


FIGURE A4 (a) Single realisation of a spatiotemporal Hawkes model (Equation 5) where the spatial self-excitation follows a bivariate Gaussian kernel with $\sigma_x = 0.7$, $\sigma_y = 0.9$ and $\rho = 0.2$. The timestamp of each point is denoted by the size of the plotting character (the larger the later the timestamp). Given an arbitrary point, X, the bivariate Gaussian kernel corresponding to the spatial self-excitation is shown by the coloured contours centred at this point. (b) The raw bias (i.e., for any parameter value θ , raw bias = $\hat{\theta} - \theta$) for each parameter of the spatiotemporal Hawkes model.